Molecular Microbiology (2003) 50(1), 167–181 doi:10.1046/j.1365-2958.2003.03697.x

Large-scale essential gene identification in *Candida albicans* and applications to antifungal drug discovery

Terry Roemer,[1*†] Bo Jiang,[1†] John Davison,[1] Troy Ketela,[1] Karynn Veillette,[1] Anouk Breton,[1] Fatou Tandia,[1] Annie Linteau,[1] Susan Sillaots,[1] Catarina Marta,[1] Nick Martel,[1] Steeve Veronneau,[1] Sebastien Lemieux,[1] Sarah Kauffman,[2] Jeff Becker,[2] Reginald Storms,[1,3] Charles Boone[1,4] and Howard Bussey[1,5]

[1]*Elitra Canada, Montreal, Quebec, Canada, H2X 3Y8 and Elitra Pharmaceuticals, San Diego, CA 92121, USA.*
[2]*Department of Microbiology, University of Tennessee, Knoxville, TN 37919, USA.*
[3]*Department of Biology, Concordia University, Montreal, Quebec, Canada, H3G 1M8.*
[4]*Banting and Best Department of Medical Research and Department of Medical Genetics and Microbiology, University of Toronto, ON, Canada M5G 1L6.*
[5]*Department of Biology, McGill University, Montreal, Quebec, Canada, H3A 1B1.*

Summary

***Candida albicans* is the primary fungal pathogen of humans. Despite the need for novel drugs to combat fungal infections [Sobel, J.D. (2000) *Clin Infectious Dis* 30: 652], antifungal drug discovery is currently limited by both the availability of suitable drug targets and assays to screen corresponding targets. A functional genomics approach based on the diploid *C. albicans* genome sequence, termed GRACE™ (gene replacement and conditional expression), was used to assess gene essentiality through a combination of gene replacement and conditional gene expression. In a systematic application of this approach, we identify 567 essential genes in *C. albicans*. Interestingly, evaluating the conditional phenotype of all identifiable *C. albicans* homologues of the *Saccharomyces cerevisiae* essential gene set [Glaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., *et al.* (2002) *Nature* 418: 387–391] by GRACE revealed only 61% to be essential in *C. albicans*, emphasizing the importance of performing such studies directly within the pathogen. Construction of this conditional mutant strain collection facilitates large-scale examination of terminal phenotypes of essential genes. This information enables preferred drug targets to be selected from the *C. albicans* essential gene set by phenotypic information derived both *in vitro*, such as cidal versus static terminal phenotypes, as well as *in vivo* through virulence studies using conditional strains in an animal model of infection. In addition, the combination of phenotypic and bioinformatic analyses further improves drug target selection from the *C. albicans* essential gene set, and their respective conditional mutant strains may be directly used as sensitive whole-cell assays for drug screening.**

Introduction

Fungal genomics couples global gene discovery with the ability to determine function through mutant phenotypes (Goffeau *et al.*, 1997; Winzeler *et al.*, 1999). Functional genomics studies in *Saccharomyces cerevisiae* have shown that once a genome is sequenced and properly annotated, the systematic deletion of essentially all individual genes is possible (Giaever *et al.*, 2002). A direct medical application of this approach involves the identification of the complete set of genes essential for the life of a fungal pathogen, which in turn define targets for drug intervention. In this view, the lethal phenotype associated with deletion of an essential gene would reflect the effect of a drug that inactivates the product of that gene. *S. cerevisiae* has often been used as a model organism for predicting targets in *Candida albicans*. However, recent studies suggest the essential role of *S. cerevisiae* genes can neither be reliably extrapolated directly to *C. albicans* (Nagahashi *et al.*, 1998; Kelly *et al.*, 2000) nor vice versa (Mio *et al.*, 1997; Lussier *et al.*, 1998). In part, this is probably because these two fungi diverged approximately 800 million years ago and therefore likely show important evolutionary divergence in their gene complements (Heckman *et al.*, 2001). In fact, using the computer algorithm BLAST at $P < 1 \times 10^{-20}$, approximately 40% of *C. albicans* genes have no identifiable *S. cerevisiae* homologue. Thus, it is perhaps not surprising, as we show here experimentally, that these two fungi differ in their capacity to buffer deleterious mutations, resulting in different sets of essential genes (see also Hartman *et al.*, 2001).

Genomic approaches performed using a non-pathogenic organism, such as *S. cerevisiae*, limit the ability to study pathogenesis and combat fungal disease.

Accepted 30 June, 2003. *For correspondence. E-mail troemer@elitra.com; Tel. (+1) 514 987 4288; Fax (+1) 514 987 0389.
†Authors have contributed equally.

Preferably, phenotypic analysis performed in *C. albicans* permits relevant conclusions to be drawn with respect to the biology of a human fungal pathogen that are applicable throughout the drug discovery process (e.g. target identification, target validation and high throughput screening). However, despite its importance in disease, *C. albicans* has proven refractile to large-scale genetic analysis for several reasons. First, this diploid organism lacks an obvious sexual cycle, precluding classic genetic analysis. Moreover until recently, researchers have lacked molecular reagents and sequence information for efficient and targeted gene deletion and for direct determination of gene essentiality (DeBacker *et al.*, 2000; Enloe *et al.*, 2000). To date, phenotypic studies have been performed on 131 *C. albicans* genes, of which 21 genes have been shown experimentally to be essential (CalPD August 19th 2002), while an antisense and promoter interference-based study identified 86 genes affecting growth (DeBacker *et al.*, 2001a). Recently, the *C. albicans* genome has been sequenced to 10.9× coverage totalling approximately 16 Mb and estimated to comprise over 9100 genes predicted to encode proteins of at least 100 amino acids or larger (www-sequence.stanford.edu/group/candida). Complete annotation of the *C. albicans* genome will provide a foundation from which systematic functional genomics strategies may be based.

The tetracycline (Tet)-regulatable promoter system is an attractive method for controlling gene expression in a variety of eukaryotic organisms including *C. albicans* (Nakayama *et al.*, 2000). Typically, this system has two components; a chimeric transactivator protein [consisting of the *Escherichia coli tet*R DNA-binding domain fused to a variety of transcriptional activation domains (Gari *et al.*, 1997)] and a tetracycline responsive promoter containing multiple Tet operator elements to which the transactivator binds. Constitutive expression of a gene under the regulation of the Tet promoter system is achieved by the binding of the transactivator protein to the Tet promoter. Conditional repression of a Tet-promoter regulated gene is achieved by providing tetracycline (and analogues thereof), to inhibit stable association between the transactivator and Tet promoter. Therefore, controlled expression of a gene under the regulation of the Tet promoter system is achieved by either constitutive expression or repression in the absence or presence of tetracycline respectively.

Here we describe a method that makes use of the *C. albicans* genome sequence and enables rapid and systematic construction of conditional mutants for large-scale genetic analysis of this fungal pathogen. The GRACE method involves two successive manipulations: (i) precise gene replacement of one allele and (ii) controllable expression of the remaining allele by replacing the native promoter with a tightly regulatable tetracycline (Tet) promoter. In this study, 1152 *C. albicans* genes have been evaluated by the GRACE method, of which, 567 genes are experimentally demonstrated as essential for growth. We also describe how the resulting conditional mutant strain collection has broad utility in providing *in vitro* and *in vivo* phenotypic information relevant to antifungal target selection and prioritization as well as providing highly sensitive whole-cell assays for drug screening.

Results

C. albicans *conditional mutant methodology*

To obtain gene function information from *C. albicans* on a genomic scale, we developed a method that allows the systematic construction of conditional mutants. A suitable parent strain (CaSS1) for our approach was engineered in the *C. albicans* CAI4 strain background (Fonzi and Irwin, 1993) by introducing a homozygous *his3* auxotrophic deletion mutation and expression of a chimeric tetracycline transactivator protein comprising the *tet*R DNA-binding domain of *E. coli* fused to the *S. cerevisiae* GAL4 activation domain (see *Experimental procedures*). The method, termed GRACE™, involves two successive steps (Fig. 1A), each involving PCR-based methodology that permits rapid construction of mutant alleles by homologous recombination (Baudin *et al.*, 1993; Wilson *et al.*, 1999). First, a precise gene replacement of one allele of the diploid pair is made in CaSS1, creating a heterozygous mutation for the gene of interest. Second, controllable expression of its remaining allele is engineered by replacing 250 bp of the native promoter upstream of the ATG codon with the tightly regulatable tetracycline (Tet) promoter. Thus, one gene copy is deleted and the other is placed under the control of a regulatable promoter. Transformants resulting from each step were genotyped by PCR methods using suitable oligonucleotide primers to verify disruption and promoter replacement junctions respectively (see *Experimental procedures*). As in rare cases, *C. albicans* genes are reported as having three allelic copies (Gow *et al.*, 1994), confirmation that an unaltered wild-type allele is absent was also performed by PCR at the conclusion of strain construction.

The GRACE™ method allows transcriptional repression of tetracycline promoter-regulated genes in two ways. First, tetracycline-dependent repression of conditionally regulated genes may be achieved by supplementing growth media with tetracycline (Gari *et al.*, 1997; Nakayama *et al.*, 2000). Alternatively, a theoretically minimal level of gene expression for tetracycline promoter-regulated genes may be attained in GRACE strains using a genetic strategy. In the GRACE system, the gene encoding the transactivator is expressed from the *URA3*-marked plasmid pRC18 that is stably integrated as a tandem duplication into one allele of the *LEU2* locus.

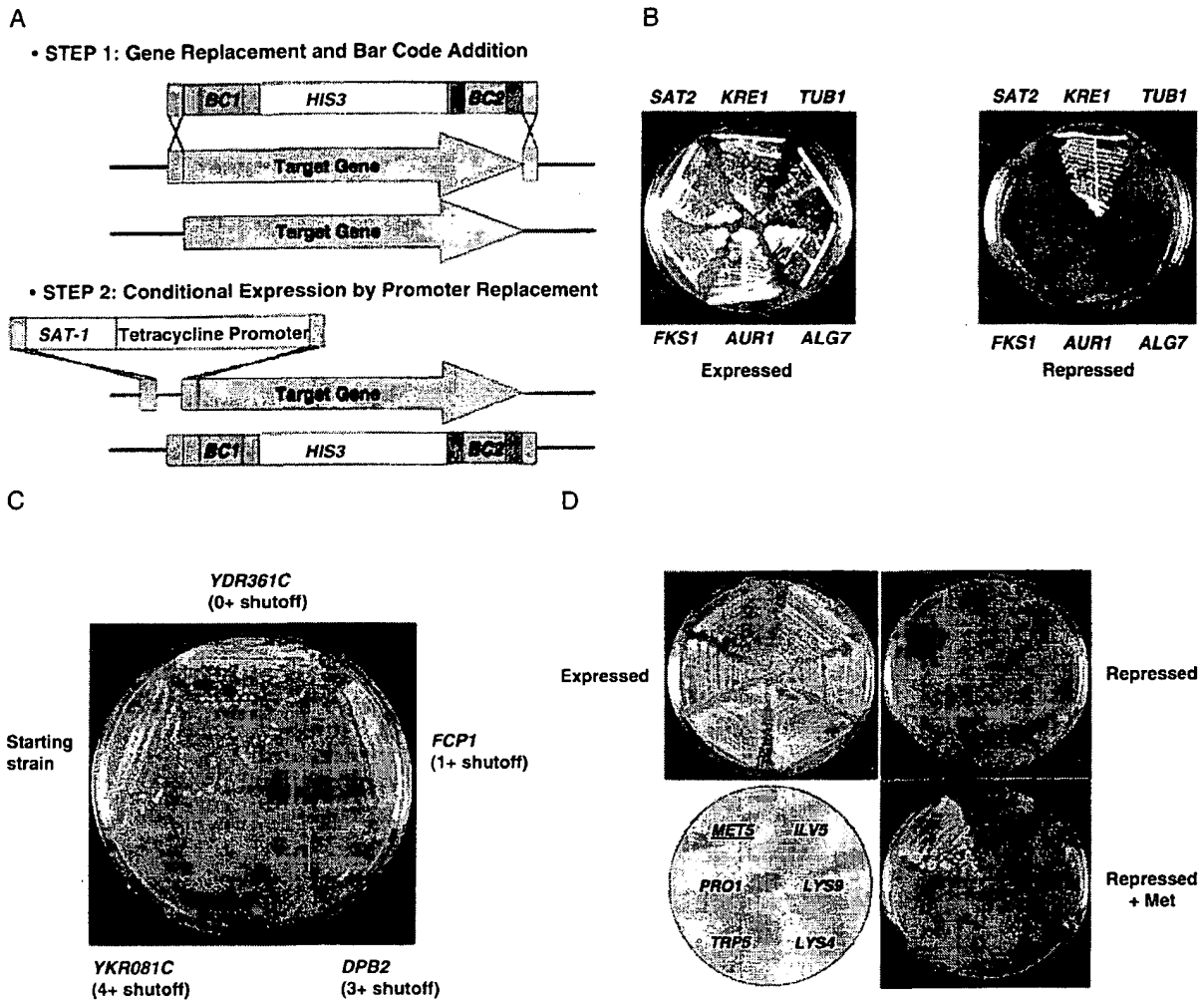

Fig. 1. The GRACE method (gene replacement and conditional expression) of target validation.

A. Step 1: heterozygote strains were constructed by transforming the wild-type *Candida albicans* starting strain, CaSS1 (see *Experimental procedures*) using a PCR-generated disruption cassette containing a *HIS3* selectable marker flanked with appropriate homologous sequence to precisely replace one allele of the target gene. Two distinct bar codes (BC1, 'up tag' and BC2, 'down tag') were introduced into the disruption cassette during PCR amplification. Two primer pairs that anneal to the common arms (yellow and red respectively) flanking each 'up tag' and 'down tag', enable simple PCR amplification of the strain-identifying bar codes. Thus, all heterozygote strains are uniquely tagged with distinct strain identifying bar codes. Step 2: bar-coded heterozygous strains were transformed using a PCR-generated tetracycline promoter replacement cassette containing the SAT-1 dominant selectable marker engineered for expression in *C. albicans*. Homologous flanking sequence was added during PCR amplification to precisely replace the endogenous promoter of the remaining wild-type allele with the Tet promoter replacement cassette after transformation.

B. Conditional mutant strains for six distinct *C. albicans* genes (*FKS1*, *TUB1*, *ALG7*, *AUR1*, *SAT2* and *KRE1*) were constructed by the GRACE method to directly evaluate gene essentiality. Note: orthologues of each of these genes (excluding *KRE1*) are reported to be essential in *Saccharomyces cerevisiae* (http://genome-www.stanford.edu/Saccharomyces/). Also, by an alternative gene knockout procedure, *FKS1* was independently confirmed to be essential for *C. albicans* growth and serves as a positive control for determining essentiality in this experiment (Douglas *et al.*, 1997). Conversely, *KRE1* has been independently disrupted by the Ura blaster method and shown to lack any detectable growth phenotype under standard laboratory conditions (data not shown), and serves as a negative control for gene essentiality by GRACE analysis. Strains were grown under non-repressing (YNB) or standard repressing conditions (YNB + 100 μg ml^{-1} tetracycline) for 2 days at 30°C and photographed.

C. Growth phenotypes of GRACE conditional mutants under standard repressing conditions are subdivided into the following categories: essential for cell viability (4+ shutoff), displaying strong growth defects (3+ shutoff), milder growth defects (2+ or 1+ depending on severity of phenotype), or no growth phenotype (0+ shutoff). The *C. albicans* strain, CaSS1, was made heterozygous for *HIS3* by homologous recombination of the *HIS3* cassette back into one *his3*-deleted allele and serves as a wild-type control for growth comparisons.

D. GRACE mutants for *MET5*, *ILV5*, *PRO1*, *LYS9*, *TRP5* and *LYS4* are grown on non-repressing (YNB), standard repressing conditions (YNB + 100 μg ml^{-1} tetracycline) or methionine-supplemented repressing conditions (YNB + 100 μg ml^{-1} tetracycline + 100 μg ml^{-1} methionine) for 2 days at 30°C and photographed. Similar results are also observed for isoleucine, proline, lysine or tryptophan-supplemented repressing conditions (data not shown).

Consequently, GRACE strains can be forced to lose the transactivator-containing plasmid by counterselection on 5-FOA containing media (Guthrie and Fink, 1991). Therefore, GRACE strains that have lost the transactivator and are unable to grow in the presence of 5-FOA-containing medium, may be independently verified as essential genes. As chromosome instability and mitotic recombination are frequently detected in *C. albicans*, growth phenotypes were examined on minimal media to select for maintenance of both the deletion allele and the transactivator.

Qualitative scores were assigned to conditional mutants to classify the severity of terminal phenotypes observed. Conditional mutants that fail to grow in the presence of tetracycline were scored as displaying a 4+ phenotype (Fig. 1B and C). Other growth phenotypes were scored as 3+ (strong), 2+ (medium), 1+ (mild) or 0+ (no growth phenotype) (Fig. 1C). We postulate that essential genes are identified in this study by mutants that produce a 3+ or 4+ growth phenotype by tetracycline repression [which includes severe growth defects potentially beyond reliable detection limits by standard methods of determining gene essentiality (Fonzi and Irwin, 1993)] or display no growth after 5-FOA counterselection. Therefore, essential genes described here are strictly those that cause dramatic or complete absence of growth under the repressing conditions examined.

As our approach to identifying essential genes relies on an efficient repression of mRNA expression from the tetracycline-regulatable promoter, reproducibility and reliability of the GRACE method was evaluated by comparing the growth phenotypes with all GRACE strains to which known phenotypes to the corresponding genes are available or are strongly suspected based on gene function. All examined genes encoding amino acid biosynthetic enzymes (*TRP5, PRO1, MET2, ILV5, LYS5,* and *LYS9*) displayed a clear 4+ conditional growth phenotype corresponding to their predicted cognate amino acid deficiency (Fig. 1D). In addition, 49 of 53 *C. albicans* genes (92%) within this collection corresponding to genes known to be essential in *S. cerevisiae* and that participate in the highly conserved essential process of protein synthesis (i.e. tRNA synthetases, translation factors, and ribosomal proteins) were experimentally shown to be essential for growth. Alternatively, a comparison between those genes represented in the GRACE collection and independently identified by an antisense and promoter interference method to uncover growth phenotypes in *C. albicans* (De Backer *et al.*, 2001a) demonstrated 13 of the 22 genes (60%) to display growth phenotypes by both methods (Table 1). However, comparison between those genes represented in the GRACE collection to which published essential or viable *C. albicans* null phenotypes have been determined ($n = 56$), a much higher level of agreement was found between the data sets for both essentiality (90%, $n = 21$) and non-essentiality (86%, $n = 35$) (Table 2). Overall reproducibility of phenotypes determined by the GRACE method could be assessed for approximately 60% of the genes examined ($n = 1152$) as two or more independent transformants after the promoter replacement step were correctly identified. A reproducible terminal phenotype was determined in over 95% of these cases. Collectively, these results demonstrate that the Tet promoter system is repressed efficiently and that this approach provides a sufficiently reliable and reproducible method for determining gene essentiality.

Table 1. Comparison between GRACE and the anti-sense method of *C. albicans* essential gene determination.

Gene	ORF6 name	Tet-off	5FOA-off	Anti-sense
ENO1	ORF6.6269	4+	Yes	Critical for growth
MPS1	ORF6.8374	4+	Yes	Critical for growth
RHO1	ORF6.4937	4+	Yes	Critical for growth
RNR1	ORF6.4851	4+	Yes	Critical for growth
RPF1	ORF6.1627	4+	Yes	Critical for growth
SFI1	ORF6.6973	4+	Yes	Critical for growth
YEF3	ORF6.5226	4+	Yes	Critical for growth
APC1	ORF6.7590	3+	No	Critical for growth
NAR1	ORF6.4598	3+	No	Critical for growth
FTH1	ORF6.3026	3.5+	No	Critical for growth
RIA1	ORF6.7264	2+	Yes	Critical for growth
RVS167	ORF6.3078	2+	Yes	Critical for growth
YAE1	ORF6.7963	2+	Yes	Critical for growth
EFB1	ORF6.7533	0+	No	Critical for growth
ENP1	ORF6.4355	0+	No	Critical for growth
FAL1	ORF6.6169	0+	No	Critical for growth
HOL1	ORF6.4874	0+	No	Critical for growth
IST2	ORF6.3721	0+	No	Critical for growth
NDE1	ORF6.7223	0+	No	Critical for growth
ORF6.6103	ORF6.6103	0+	No	Critical for growth
ORF6.8597	ORF6.8597	0+	No	Critical for growth
SHA3	ORF6.6480	0+	No	Critical for growth

Large-scale essential gene identification

To evaluate the concordance of gene essentiality between *S. cerevisiae* and *C. albicans*, a large-scale genetic analysis of *C. albicans* genes sharing homology to known *S. cerevisiae* essential genes was performed. Of the estimated 6281 *S. cerevisiae* genes, 1105 are essential for growth at 30°C (Giaever *et al.*, 2002). Our analysis of the 9168 possible *C. albicans* open reading frames (ORFs) annotated by the Stanford Genome Centre, genome revealed a total of 6650 non-allelic and non-overlapping ORFs encoding proteins of 100 amino acids or larger (see *Experimental procedures*). A very similar *C. albicans* gene number has been recently reported (DeBacker *et al.*, 2001b). A total of 864 *C. albicans* genes sharing significant homology (BLAST $P < 1 \times 10^{-10}$) to *S. cerevisiae* essential genes were identified; of which, GRACE strains for 823 genes (or 95%) could be constructed. Of these,

Table 2. Concordance between GRACE and current methods of *Candida albicans* essential gene determination.

Essential by both methods					Non-essential by both methods				
Gene	ORF6 name	Tet-off	5FOA-off	Literature (CalPD)	Gene	ORF6 name	Tet-off	5FOA-off	Literature (CalPD)
ARG5,6[a]	ORF6.4898	4+	Yes	Auxotrophic	*ALK8*	ORF6 1706	0+	No	Viable
CCT8	ORF6.6242	4+	Yes	Lethal	*ALT2*	ORF6.3984	0+	No	Viable
CDC28	ORF6.5246	3+	Yes	Lethal	*CAG1*	ORF6.7428	2+	No	Viable
CET1	ORF6.4276	3+	No	Lethal	*CDC25*	ORF6.7312	0+	No	Viable
CHS1	ORF6.3499	3.5+	Yes	Lethal	*CDC43*	ORF6.1576	2+	No	Viable
DPB2	ORF6.8912	3+	No	Lethal	*CHK2*	ORF6.7281	0+	No	Viable
FKS1	ORF6.3785	4+	Yes	Lethal	*CHT2*	ORF6.2344	2+	No	Viable
GNA1[a]	ORF6.3848	4+	Yes	Auxotrophic	*CNH1*	ORF6.2059	0+	No	Viable
GSP1	ORF6.4664	4+	Yes	Lethal	*CRK1*	ORF6.308	0+	No	Viable
HEM3[a]	ORF6.1095	3+	No	Auxotrophic	*CYR1*	ORF6.5946	0+	Yes	Viable
HIS1[a]	ORF6.7439	4+	Yes	Auxotrophic	*ERG24*	ORF6.1086	1+	No	Viable
KRE9[a]	ORF6.5053	4+	Yes	Conditional lethal	*ERG6*	ORF6.6441	0+	No	Viable
NMT1	ORF6.2367	3+	No	Lethal	*GLY1*	ORF6.5812	0+	No	Viable
PRS1	ORF6.8218	4+	Yes	Lethal	*HOG1*	ORF6.7282	0+	No	Viable
PSA1	ORF6.5411	4+	Yes	Lethal	*ICL1*	ORF6.6993	1+	No	Viable
SEC14	ORF6.8190	3+	No	Lethal	*KEX2*	ORF6.4600	2+	No	Viable
SEC20	ORF6.4468	4+	No	Lethal	*MNT2*	ORF6.5738	0+	No	Viable
SEC4	ORF6.4672	4+	Yes	Lethal	*NAG2*	ORF6.2471	0+	No	Viable
YPT1	ORF6.7789	4+	Yes	Lethal	*NIK1*	ORF6.7281	0+	No	Viable
					NTC1	ORF6.8731	2+	No	Viable
Essential by GRACE but non-essential by current methods					*PKC1*	ORF6.9136	1+	No	Viable
					PLB99	ORF6.3690	0+	Yes	Viable
MIG1	ORF6.3170	3+	No	Viable	*SKN1*	ORF6.8523	2+	No	Viable
MNN9	ORF6.8543	3+	Yes	Viable	*SLN1*	ORF6.8256	1+	No	Viable
PHR2[a]	ORF6.6260	3.5+	Yes	Viable	*SSK1*	ORF6.7978	0+	No	Viable
PRA1	ORF6.6934	3.5+	No	Viable	*SUC1*	ORF6.8400	0+	No	Viable
RIM8	ORF6.6250	3+	No	Viable	*TPK2*	ORF6.4228	2+	No	Viable
					TPS1	ORF6.6537	0+	No	Viable
Non-essential by GRACE but essential by current methods					*TPS2*	ORF6.7775	0+	No	Viable
ESS1	ORF6.3412	2+	No	Lethal	*XOG1*	ORF6.1982	1+	No	Viable
KRE6	ORF6.8523	2+	No	Lethal					

a. Conditional essential phenotype.
All references may be obtained through http://www.proteome.com.

506 (61%) showed a dramatic growth defect with a score of 3+ or higher. Further, a reciprocal analysis was performed by examining the growth phenotype of 258 *C. albicans* genes sharing homology to genes known to be non-essential for growth in *S. cerevisiae*. Over 20% of this gene set was scored as essential for growth in *C. albicans*. To examine conservation of gene essentiality over two more distantly related yeasts, an additional 27 *C. albicans* genes whose homologue is reported to be essential in *Schizosaccharomyces pombe* but non-essential in *S. cerevisiae* were examined, revealing nine homologues (33%) in *C. albicans* scoring a 3+ or higher growth phenotype. Finally, we randomly picked 44 genes, none of which had an essential counterpart in *S. cerevisiae* or *S. pombe*, and found that seven (16%) scored 3+ or higher growth phenotype. Summing the 506 genes known to be essential in *S. cerevisiae* and demonstrated to be essential in *C. albicans* by GRACE, plus a prediction that 16% of the remaining 5840 C. albicans genes are probably essential, we estimate approximately 1400 (21%) *C. albicans* genes are required for growth, a proportion quite similar to the proportion (18%) of essential genes found for *S. cerevisiae* (Winzeler *et al.*, 1999; Giaever *et al.*, 2002). These results suggest that significant non-overlap exists between the essential gene sets of *C. albicans* and *S. cerevisiae*, despite the genomes of both yeasts sharing a similar percentage of essential genes.

Essential gene identification in *C. albicans* has been extended beyond those genes known to be essential in *S. cerevisiae* or *S. pombe*. To date 1152 *C. albicans* genes have been examined, of which, 567 were shown to be essential. To view the distribution of each growth inhibition phenotype (4+ to 0+) by functional classification, we assigned a functional annotation to the *C. albicans* genes corresponding to the 15 distinct functional categories described by MIPS for their *S. cerevisiae* homologues (http://www.mips.gsf.de) (Fig. 2). An additional 16th category was created for *C. albicans* genes with no *S. cerevisiae* homologue (Blast *E*-value $P > 1 \times 10^{-20}$) currently representing 8% of the essential gene set. The *C. albicans* essential gene set encodes proteins involved in a broad spectrum of basic cellular processes, as well as a substantial subset (20%) whose biological function remains unknown in either *S. cerevisiae* or *C. albicans*. Interest-

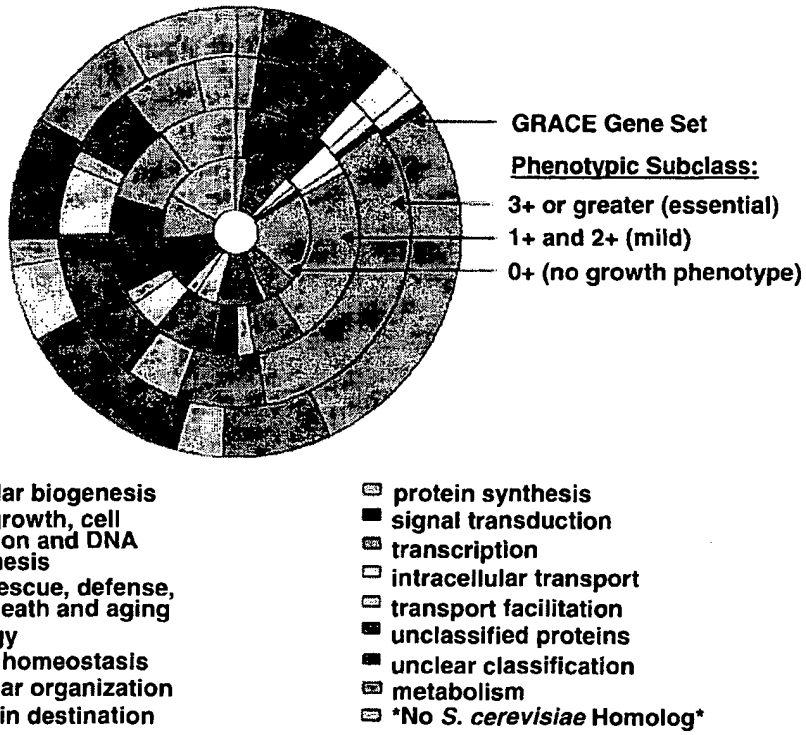

Fig. 2. Summary of the GRACE strain collection according to phenotype and functional annotation. A pie chart containing four concentric rings illustrates the relative abundance of genes within each phenotypic subclass predicted to participate in 15 basic cellular processes annotated for the *S. cerevisiae* homologue by MIPS (http://www.mips.biochem.mpg.de). The outermost ring summarizes the relative abundance of identified genes within the GRACE 1152 gene set predicted to participate either in any of these 15 basic cellular processes or annotated as lacking a clear *S. cerevisiae* homologue using a Blast *E*-value cutoff of 1×10^{-20}. Internal concentric rings summarize the functional classification and relative abundance of the essential gene subset and nonessential gene subsets (i.e. 0+ and 2+ growth phenotypes).

ingly, the largest overlapping essential gene set between *S. cerevisiae* and *C. albicans* was functionally annotated to the category of cellular organization. This may reflect a high degree of conservation between pathways controlling processes such as cytoskeletal organization and polarized morphogenesis in both organisms.

Comparative analyses to prioritize essential genes as antifungal drug targets

Historically antifungal drug discovery has been restricted to a limited set of validated targets, with fewer than 10 genes encoding products that are targeted by current antifungal drugs (DiDomenico, 1999). The approach described here makes it possible to identify a comprehensive target set, amongst which, a preferred subset of suitable drug targets may be selected. For example, essential fungal proteins absent from the human proteome may represent ideal targets because their inhibitors may have less toxic side-effects in humans. Moreover, provided these targets are also conserved amongst major fungal pathogens, their inhibitors could serve as broad-spectrum antifungal agents. To facilitate this type of target prioritization, a 'target space' plot displaying the similarity of all *C. albicans* genes against *S. cerevisiae* and human genes and with the GRACE gene set highlighted, is shown (see Fig. 3A). Globally displaying genes by their relative relatedness illustrates that *C. albicans* genes are typically more closely related to *S. cerevisiae* genes than to their human counterpart(s). An interesting exception is the *C. albicans* essential gene, ORF6.7629 (ftp:/cycle.stanford.edu/pub/projects/candida/, see Fig. 3B), which encodes a NADH-Ubiquinone oxidoreductase subunit that is highly conserved in humans but absent in *S. cerevisiae* (Malpertuy *et al.*, 2000). Applying this map to continued target discovery enables a rational and systematic approach to prioritizing all potential targets that remain to be evaluated by GRACE analysis.

To summarize the currently identified *C. albicans* target space, a second plot restricted to the *C. albicans* GRACE set is shown highlighting genes essential for growth (Fig. 3B). Known antifungal targets including *EFT2, ERG11, ERG1, FKS1*, which span a broad range of fungal and human relatedness, are included to emphasize the diversity of demonstrated 'drug-able' targets previously exploited. Our phenotypic annotation of this target space identifies numerous *C. albicans* genes strongly conserved with *S. cerevisiae* and absent from humans (upper left corner of the plot), which may represent broad spectrum targets. Essential genes located in the lower left corner of the plot (Blast *E*-value $P > 1 \times 10^{-20}$), which are absent from both *S. cerevisiae* and human genomes, represent nearly 6% ($n = 33$) of the current essential gene set. BLAST analysis of this gene set against both *S. pombe* (http://www.sanger.ac.uk/Projects/S_pombe/) and *N. crassa* (www-genome.wi.mit.edu/annotation/fungi/neurospora/) genomes identified 27 of the 33 genes as absent from both genomes (Blast *E*-value $P > 1 \times 10^{-20}$) and are therefore suspected '*C. albicans*-specific' genes potentially suitable as narrow spectrum *C. albicans* drug targets. Similar genomic comparisons could be made with whole genome sequence information from additional human fungal pathogens, such as *Aspergillus fumigatus* and *Cryptococcus neoformans*, which would accurately establish the identity of both broad and narrow spectrum target sets within the *C. albicans* essential gene set. Towards this goal, we have recently completed the first detailed sequence coverage of *A. fumigatus* in collaboration with Celera Genomics (B. Jiang, S. Lemieux, W. Hu and H. Bussey, in prep.).

Phenotypic analyses for prioritizing antifungal drug targets

Another important criterion to prioritize potential drug targets amongst this target set is to require that the gene not only be essential for growth, but that its inhibition also leads to actual death of the pathogen. In this way essential targets can be classified as 'cidal', in contrast to those that prevent growth or cell division and are 'static'. In critically ill systemically infected patients, drug intervention acting on cidal targets should allow a rapid clearance of the fungal infection and provide a major clinical advantage. Examination of the essential genes associated with a 4+ growth inhibition phenotype ($n = 403$) revealed that at least 20% ($n = 80$) displayed a rapid cidal terminal phenotype, as scored by loss of colony forming ability after 20 min of exposure to tetracycline. An additional 33% ($n = 132$) of *C. albicans* essential genes examined displayed a cidal terminal phenotype detected over a 2 day period. Therefore, as in *S. cerevisiae* (Buurman *et al.*, 2001), a substantial proportion of *C. albicans* essential genes are required for both cell growth and viability.

Definitive target validation requires experimental demonstration that a gene participates in a disease process, which involves an analysis of whether the gene is essential for growth within a host animal, a situation more closely resembling a human infection than *in vitro* tests. We have examined the consequences of transcriptionally repressing *C. albicans* target genes in an *in vivo* mouse infection model, where mice were administered tetracycline either prior to infection, or after, an established infection (Fig. 4). In three separate experiments, mice were infected with 10^6 cells of the corresponding GRACE strains for three targets of varying terminal phenotypes [*RHO1* (4+, rapid cidal), *ALG7* (4+, cidal) and *YBR070c* (3+)]. All mice receiving tetracycline treatment prior to infection to repress essential gene expression from the

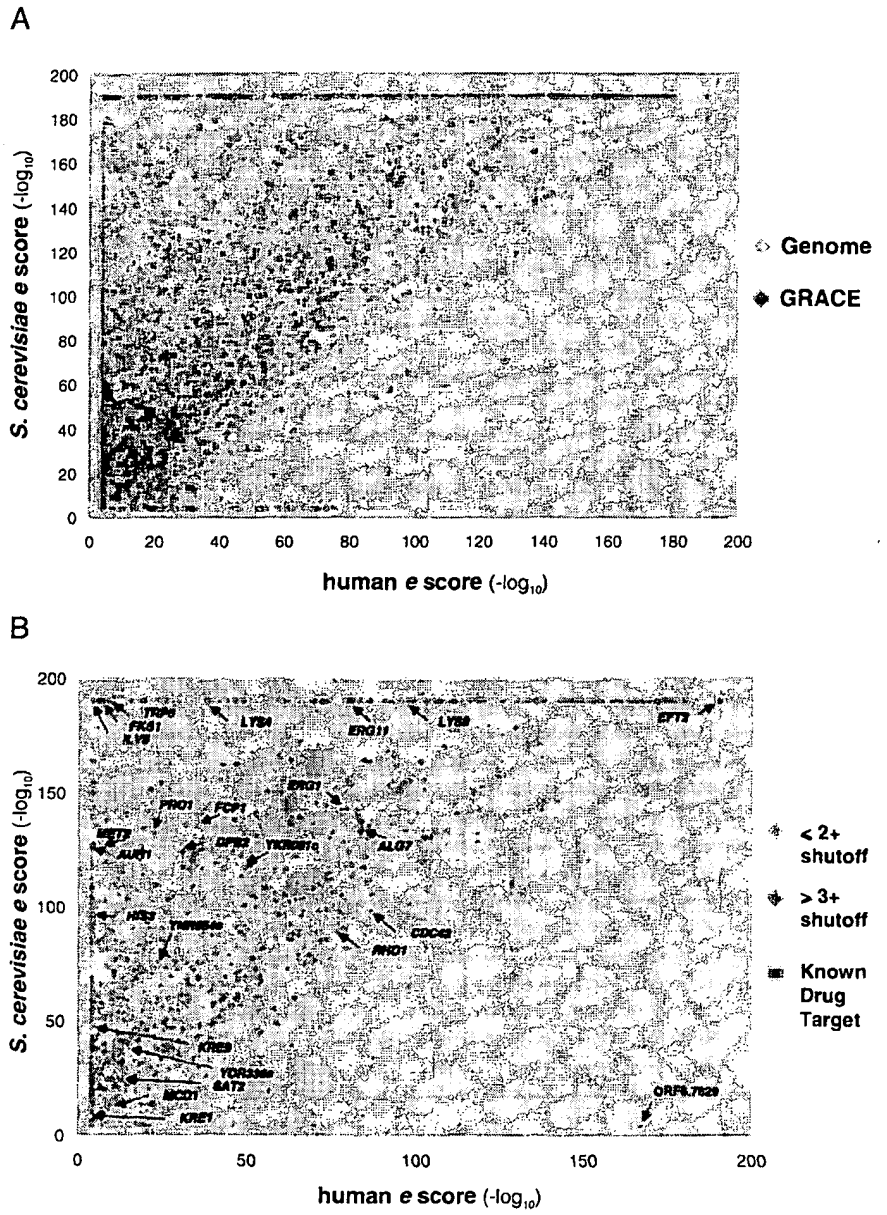

Fig. 3. Scatterplot analyses of *C. albicans* predicted proteome and GRACE target set against *S. cerevisiae* and humans. *C. albicans* predicted protein sequences (n = 6650) were compared against both *S. cerevisiae* and the human annotated protein databases by BLAST analysis to illustrate the global *C. albicans* target space and its relative conservation in *S. cerevisiae* and humans (A). Genes displaying a high degree of conservation to *S. cerevisiae* and are either absent or revealing limited conservation to humans are potential broad-spectrum antifungal drug targets (upper left corner of plot). Potential narrow spectrum (i.e. *C. albicans*-specific) putative drug targets cluster around the bottom left corner of the plot. (B) Scatterplot illustrating the target space occupied by the *C. albicans* 567 essential gene set when similarly compared against *S. cerevisiae* and humans. For reference, several known drug targets, including *ERG11*, *ERG1*, *FKS1* and *EFT2*, for which antifungal therapeutic agents have been developed (azoles, terbinafin, caspofungin and sordarins respectively) are highlighted.

onset of infection survived and remained healthy, whereas all control mice succumbed to infection within 7–10 days. These results indicate that each GRACE strain is normally virulent and that all three targets are necessary for the establishment of a fungal infection. Importantly, the requirement of *YBR070c* for *in vivo* growth and virulence in an animal model of infection and that *YBR070c* displays a strong *in vitro* growth phenotype (3+ shutoff) supports its inclusion (as well as other genes displaying similar 3+ shutoff phenotypes) as a member of

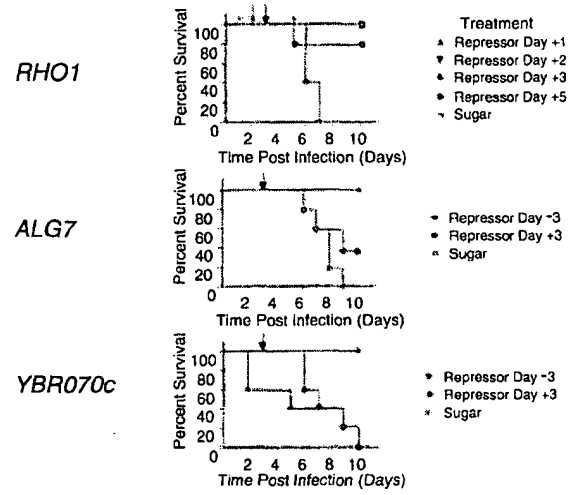

Fig. 4. Target validation and prioritization of *C. albicans* genes in an immunocompetent mouse model of systemic infection. ICR male mice were injected into the tail vein with 1×10^6 cells of one of three GRACE strains shown to display differing terminal phenotypes *in vitro*: *YBR070c* (a 3+ growth phenotype), *ALG7* (an essential gene displaying a cidal terminal phenotype) and *RHO1* (an essential gene displaying a rapid cidal terminal phenotype). Mice were given either sucrose (5%) or sucrose (5%) and tetracycline (2 mg ml^{-1}) in their drinking water starting 3 days prior to the *C. albicans* inoculation and maintained under this regimen for a period of 11 days post-infection. Inactivating expression of *YBR070c*, *ALG7*, or *RHO1* during an established infection was achieved by adding tetracycline (2 mg ml^{-1}) to the sucrose (5%) supplemented drinking water at the various post-infection time points as indicated.

the *C. albicans* essential gene set. As this regimen mimics a prophylactic setting, each of the targets examined are thus validated as drug targets suitable for prophylactic drug intervention.

In an alternate regimen that models treatment of an established systemic infection, groups of mice were infected with each of the three *C. albicans* GRACE strains and administered tetracycline to repress gene expression at various times postinfection (Fig. 4). Here, inactivating *YBR070c* at 3 days post-infection provided no beneficial effect and repression of *ALG7* resulted in only a partial extension of animal survival. In contrast, repressing *RHO1* at 1, 2 or 3 days post-infection rescued all mice, which otherwise died by 7 days when tetracycline was not administered. Furthermore, repression of this target as late as 5 days post-infection, rescued all remaining mice. Therefore, regulatable expression offered by GRACE strains can be extended to an established systemic infection model *in vivo*, enabling genes, such as *RHO1*, to be validated as therapeutic targets for treating a systemic *C. albicans* infection. Extending such genetic analyses will facilitate definitive identification and prioritization of numerous novel prophylactic and systemic infection drug targets.

Sensitized whole cell assays for drug screening

Traditionally, drug screening has been limited to those targets for which a biochemical assay could be established. Titratable repression of the Tet promoter by the addition of tetracycline allows control over the cellular concentration of a specific target protein. In most instances, this provides a mechanism of action-based assay, as cells with reduced target protein levels become hypersensitive to compounds that inhibit target compared with wild-type cells (Giaever *et al.*, 1999; DeBacker *et al.*, 2001a; Forsyth *et al.*, 2002). This is illustrated in Fig. 5, where the relative compound-specific sensitivity of three strains affecting the targets of fluconazole (Erg11p), tunicamycin (Alg7p) and 3-amino triazole (His3p) are shown. In each case, the sensitizing effect is specific to the strain depleted of the compound's cognate target; for example, reducing the level of Erg11p hypersensitizes cells to fluconazole 100-fold, while resulting in no increased sensitivity to Amphotericin B, tunicamycin or 3-amino triazole, which have distinct targets. In principle, this whole cell assay approach enables screening any target within the target space of essential genes irrespective of its biochemical function.

Discussion

This work significantly expands upon existing methodologies for determining gene function in *C. albicans*, and is particularly suited towards studying essential genes both *in vitro* and *in vivo*. We demonstrate the utility of GRACE methodology in a systematic large-scale genetic analysis in *C. albicans*. To date, we have identified and prioritized 567 essential genes as potential drug targets by using comparative analyses with *S. cerevisiae* and humans to evaluate target spectrum and specificity parameters. Phenotypic analyses identify essential genes displaying cidal terminal phenotypes and whether their genetic inactivation in a systemic model of candidiasis improves animal survival. As target inactivation studies performed *in vivo* provide a genetic prediction that chemical inhibitors to such gene products may similarly display a strong therapeutic effect, such gene products represent validated targets for antifungal drug discovery.

Construction of conditional mutants overcomes two principal technical limitations encountered with the popular Ura Blaster method of gene disruption typically used in *C. albicans*; it enables direct demonstration of gene essentiality rather than statistical inference the gene is essential based on the inability to identify viable homozygous deletion mutants and it eliminates the cumbersome steps required to recycle the single *URA3* marker (Fonzi and Irwin, 1993). Recently, an alternative strategy enabling demonstration of gene essentiality in *C. albicans*

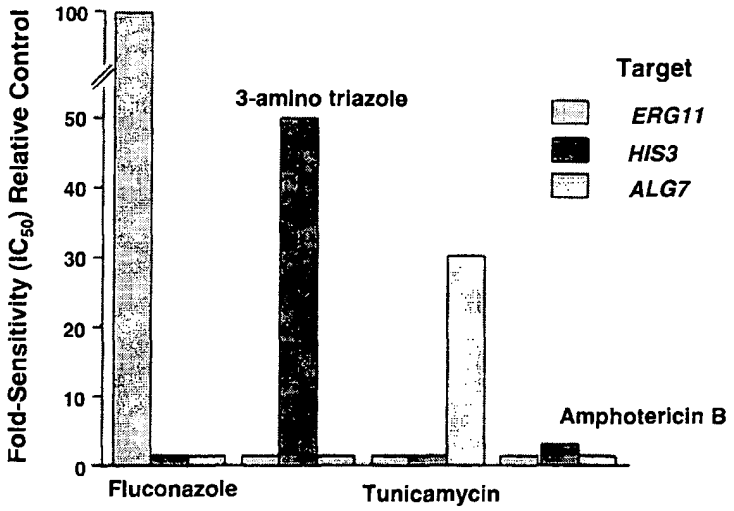

Fig. 5. *C. albicans* sensitized whole cell assays. GRACE strains conditionally regulating the known drug targets, *ERG11*, *HIS3* and *ALG7* were constructed and IC_{50} values were determined against a matrix of antifungal compounds including their cognate inhibitors, fluconazole, 3-amino triazole and tunicamycin respectively. All strains were assayed in a suitable tetracycline concentration to underexpress the drug target to such an extent where growth rate is reduced ~90%. As indicated, sensitized cells display a range between 30-fold and 100-fold lower IC_{50} values that are specifically detected between the drug target and its known inhibitor. IC_{50} determination to Amphotericin B, whose mechanism of action is distinct from fluconazole, 3-amino triazole and tunicamycin and involves disrupting the plasma membrane, revealed no elevated drug sensitivity amongst any of the three sensitized *C. albicans* strains.

has been reported (Enloe *et al.*, 2000). Termed the HT test, this method utilizes a one-step transformation and homozygosis of the deleted allele, followed by selection of a recombined second marker and examination of allelic triplication. Hence marker recycling and disruption of the second allele are elegantly achieved. For non-essential genes, PCR tests of the resulting segregants demonstrate a wild-type allele which may or may not exist with similar frequency. Alternatively, segregants disrupted of an essential gene are identified as faithfully carrying an allelic triplication of the wild-type allele. However, extensive genetic analyses are required to verify that all segregants carry allelic triplications and that potential issues of homozygosis have not complicated interpretation of the deduced essential phenotype (Enloe *et al.*, 2000). Alternatively, the conditional mutation strategy outlined here enables gene essentiality to be determined directly and the resulting laboratory strains provide a permanent resource to fully examine the range of terminal phenotypes associated with loss of any essential gene function. The titratable repression of the tetracycline promoter also enables intermediate phenotypes (i.e. hypomorphs) for further analysis of essential genes.

Although individually, a PCR-based gene disruption method (Wilson *et al.*, 1999), an alternative dominant selectable marker (Beckerman *et al.*, 2001) and a related Tet promoter system (Nakayama *et al.*, 2000) have each been previously applied to *C. albicans*, we have combined these elements and introduced additional modifications to improve their utility. Construction of both a SAT-1 dominant selectable marker system suitable for *C. albicans* as well as a *HIS3* auxotrophic marker enables conditional mutant strains to be constructed rapidly as recycling of the Ura marker is avoided, unlike that of previously reported promoter replacement-based strategies (Nakayama *et al.*, 2000; Warit *et al.*, 2000). PCR-based gene disruption strategies may now be extended to prototrophic clinical isolates using the SAT-1 marker in combination with the MPA resistance marker (Beckerman *et al.*, 2001) to completely delete each allele of a gene without recycling markers. Linking the Tet transactivator protein to a stably integrated *URA3*-plasmid also enables selection for the loss of the transactivator-containing plasmid on 5-FOA-containing medium. This provides an independent genetic test that ensures the theoretical basal level of expression from the Tet promoter is achieved by eliminating the transactivator protein from the cell.

A number of phenotypic tests were performed to extensively demonstrate both the reproducibility and reliability of the GRACE method for determining gene essentiality. Clear conditional essential phenotypes were observed for all genes involved in amino acid biosynthesis (Fig. 1D). Moreover, GRACE analysis of all genes involved in protein synthesis ($n = 53$) yielded confirmed essentiality in all but four cases. A significant overlap was also observed in the growth phenotypes of 22 genes evaluated by GRACE (Table 1) and also identified by a novel antisense and promoter interference strategy (DeBacker *et al.*, 2001a). Although in the majority of cases, genes identified as critical for growth by antisense and promoter interference strategy were either essential or shared strong growth phenotypes as conditional mutants, a significant number (40%) displayed no observable growth phenotype by GRACE on either tetracycline or 5-FOA-containing media. As this may reflect possible deleterious effects caused by antisense overexpression, examination of all genes evaluated by the GRACE method that have been independently disrupted using well-accepted genetic gene disruption methods was also performed (Table 2). Here, a high concordance in essential (90%) and non-essential

(86%) gene determination was observed, further validating the application of the GRACE procedure to a systematic and large-scale *C. albicans* essential gene identification strategy.

Minimal disagreement between the essential phenotypes determined by GRACE and Ura Blaster methods were observed. In part, this probably reflects that the basal level of expression for some genes under the control of the Tet promoter under repressing conditions (i.e. tetracycline and/or 5-FOA counterselection) remains sufficient for cell growth. This limitation to the tetracycline (or any) conditional promoter strategy would probably contribute to the failure to demonstrate a complete concordance in the phenotypes determined by gene disruption methods. However, the *C. albicans* Ura Blaster gene disruption method is particularly limited in differentiating essential genes from those displaying strong growth phenotypes. Therefore, in cases where the GRACE method identified substantial growth phenotypes (e.g. *KRE6* and *ESS1*) but these genes are determined to be essential by the Ura Blaster method (Mio *et al.*, 1997; Devasahayam *et al.*, 2002), it is unclear whether this disagreement results from our inability to sufficiently repress these genes or whether the slow growth phenotypes we detect may have obscured the identification of viable homozygotes by the Ura Blaster method. It is also possible that replacing the endogenous promoter with any alternative promoter may result in constitutive expression of genes regulated in a cell cycle-dependent manner and/or causing elevated gene expression versus wild-type cells. As such, this could result in the corresponding GRACE strain displaying *in vitro* growth and/or *in vivo* avirulence phenotypes even in the absence of tetracycline or 5-FOA. Despite these concerns, it has not precluded our ability to construct GRACE strains for ~95% of the genes we have examined. Moreover, *in vitro* growth phenotypes are only rarely seen without tetracycline and *in vivo* studies of over 100 targets (data not shown) has yet to uncover GRACE strains displaying any obvious avirulence phenotype in a sugar control group of mice. Nonetheless, such limitations may further contribute to any of the five GRACE strains in Table 2 where strong growth or essential phenotypes are detected but viable homozygous gene deletion mutants are reported. Other differences in terminal phenotypes observed may be attributable to different growth conditions used to score phenotypes (e.g. media, pH and temperature) or non-isogenic strain backgrounds between *C. albicans* isolates known to arise from chromosomal rearrangements or duplications (Magee, 1993; Rustchenko *et al.*, 1994). Construction of a plurality of Tet-promoter cassettes and/or expression levels of the transactivator to vary constitutive and basal levels of expression covering a wider dynamic range of expressed proteins in *C. albicans* may further assist genetic analyses.

We demonstrate that a principal advantage to our approach is its large-scale application in a human fungal pathogen to directly and systematically identify an unprecedented collection of *C. albicans* genes essential for growth. Our work demonstrates only approximately 61% of those *C. albicans* genes sharing homology to known *S. cerevisiae* essential genes essential when examined by GRACE. Similarly, approximately 21% of genes that display significant homology (Blast *E*-value $P < 1 \times 10^{-20}$) to a unique non-essential homologue in *S. cerevisiae*, are nonetheless essential genes in *C. albicans*. Therefore, notwithstanding our estimate that the percentage of essential genes representing the genomes of *C. albicans* and *S. cerevisiae* are quite similar (21% and 18% respectively), we identify considerable non-conservation between the essentiality of homologous genes common to these two yeasts. Although this estimate may be somewhat lower than the true essential gene overlap (due to technical limitations outlined above), the basis of this high non-concordance in gene essentiality probably relates to biological and developmental distinctions between these fungal organisms. In part, the considerable non-overlap in *S. cerevisiae* and *C. albicans* genomes, in which nearly 40% of *C. albicans* genes lack a clear homologue in yeast, may contribute to 'buffering' many mutations lethal for *S. cerevisiae*. Also, the broad distribution in homology at the primary amino acid sequence level between identifiable *S. cerevisiae* and *C. albicans* homologues suggests that in many cases homologous genes may have evolved distinctions in their specific function and/or regulation. It is also likely that unique aspects of *S. cerevisiae* and *C. albicans* fungal lifestyles, namely growth on a grape versus a pathogenic existence, probably contribute to significant non-overlapping essential gene sets. Just as a reliance on experimental approaches amenable to demonstrating similarities between these yeasts (e.g. *C. albicans* genomic library screens and forward genetics strategies based on *S. cerevisiae*) have entrenched our understanding of the relatedness between these yeasts, completion of the *C. albicans* genome and accompanying functional genomics approaches are likely to reveal more their differences.

The essential genes identified in this study serve as a broad collection of novel antifungal drug targets. This target set, combined with comparative genomics, allows rational target prioritization by both *in vitro* examination of cidal versus static terminal phenotypes and definitive target validation within an animal model of infection. In this way it is possible to not only identify novel *C. albicans* drug targets absent in *S. cerevisiae*, but also to triage amongst the known *S. cerevisiae* essential gene set of over 1100 genes (Giaever *et al.*, 2002), that subset of essential *C. albicans* orthologues which are experimentally demonstrated to meet rigorous *in vitro* and *in vivo* criteria as

preferred drug targets. The GRACE strains also serve as sensitized whole cell assays suitable for high throughput drug screening a wide spectrum of preferred targets regardless of whether functional or biochemical information is available. Furthermore, the conditional mutant strain collection overcomes the obstacles of performing genetic screens in *C. albicans* and permits rapid phenotypic screening with defined mutants to identify genes involved in a wide range of cellular processes, including growth under conditions of hyphal development (Madhani and Fink, 1998; Mitchell, 1998; Odds *et al.*, 2001), biofilm formation (Chandra *et al.*, 2001) and *in vivo* screens to identify virulence factors (Madhani and Fink, 1998; Mitchell, 1998; Brown *et al.*, 2000). The conditional mutant collection also provides a rich resource from which transcriptional profiling may be applied to essential drug targets (Hughes *et al.*, 2000). Constructing a comprehensive *C. albicans* conditional mutant strain collection should significantly accelerate functional annotation of the *C. albicans* genome to aid both basic research and antifungal drug discovery.

Experimental procedures

CaSS1 strain construction

The GRACE method involves the following steps. A complete deletion of the *HIS3* gene was constructed in CAI4 using the Ura Blaster method (Fonzi and Irwin, 1993). Briefly, a *his3* disruption cassette containing the *hisG-URA3-hisG* cassette and replacing the entire *HIS3* open reading frame (ORF) including an additional 160 bp of upstream and 200 bp downstream UTR sequence was constructed using standard molecular methods (Sambrook *et al.*, 1989). Deletion of additional 5′ and 3′ *HIS3* UTR sequence eliminates all homologous sequence between *his3::HISG* deletion alleles and the *HIS3* selectable marker of the gene disruption cassette used during Step 1 of the GRACE method. The resulting CAI4-based *ura3, his3* auxotrophic strain was confirmed by Southern blot and PCR analysis (data not shown). Construction of the *his3* null mutation deletes 200 bp of possible promoter sequence of a predicted downstream ORF, PET56, which is required for mitochondrial respiration. PET56 gene function has not been significantly affected during construction of the *his3* auxotrophy as CaSS1 grows normally on 2% glycerol, a non-fermentable carbon source.

Tetracycline-dependent transactivation fusion proteins were constructed by fusing in frame the *E. coli tet*R tetracycline repressor domain (aa 1–207) to the activation domain of either the *S. cerevisiae GAL4* (aa 785–881) or *HAP4* (aa 424–554) into the *C. albicans* plasmid, pRC18 (Stoldt *et al.*, 1997). Multiple CTG codon corrections were introduced to comply with the *C. albicans* genetic code (Santos and Tuite, 1995) using the Quikchange Site-Directed Mutagenesis kit (Stratagene). Constitutive expression of the resulting fusion constructs was achieved using *ACT1* promoter and *ACT1* terminator sequences. pRC18-based *tet*R-Gal4AD or *tet*R-Hap4AD transactivator plasmids were linearized by *Kpn*I digestion, transformed using standard transformation methods (Braun and Johnson, 1997) and Ura3 + transformants were selected on YNB medium containing 100 μg ml^{-1} histidine. Linearized plasmid prior to transformation directs stable plasmid integration into one allele of *LEU2* resulting in tandem duplication in one allele (creating both 5′ and 3′ truncation copies of *leu2*) after transformation. The resulting *his3::hisG/his3::hisG* auxotropic strain constitutively expressing the *tet*R-GAL4AD (and Ura3+) was confirmed by Southern blot analysis and named CaSS1. Heterozygote barcoding required the construction of a *HIS3* gene disruption cassette. A functional *HIS3* gene fragment containing sufficient promoter and UTR sequence was subcloned into Bluescript plasmid, PBSK (Stratagene) and serves as an auxotrophic marker. *HIS3*-based gene disruption cassettes were constructed by PCR amplification from PBSK-*HIS3* using 99 mers containing (i) 43 bp of flanking sequence identical to the sequence 5′ or 3′ of the *C. albicans* target gene and (ii) internal strain-identifying 56 bp barcode cassettes (Shoemaker *et al.*, 1996; Giaever *et al.*, 1999; Giaever *et al.*, 2002) containing unique 20 bp barcodes and flanked by common arms, 18 bp of which at the 3′ end of the oligonucleotide are complementary to the PBSK-*HIS3* cassette. *C. albicans* heterozygote bar-coded gene disruptions were obtained after CaSS1 transformation with the PCR-amplified *HIS3* gene disruption cassette and selection for *HIS3* prototrophs on YNB medium. PCR analysis was performed to verify correct integration. This process was repeated for all heterozygote barcode strains comprising the conditional expression strain set presented. The SAT-1 marked Tet promoter replacement cassette construction first involved construction of a dominant selectable marker for *C. albicans*. *C. albicans* is sensitive to the nucleoside-like antibiotic nourseothricin at a concentration of 200 μg ml^{-1} (data not shown). A dominant selectable marker conferring nourseothricin drug resistance was constructed using the *E. coli* SAT-1 gene. Heterologous expression of the SAT-1 gene in *C. albicans* involved mutagenesis of a single CTG codon in the SAT-1 ORF to a CTT codon complying with the *C. albicans* genetic code and by introducing a *ACT1* promoter and a *PCK1* terminator sequence. We refer to this dominant selectable marker as SAT-1. A SAT-1-marked Tet promoter replacement cassette was constructed by fusing the *S. cerevisiae*-based tetracycline promoter derived from pCM184 (Gari *et al.*, 1997) directly after the *PCK1* 3′ UTR of SAT-1. The tetracycline promoter system contains an *ADH1* 3′ terminator sequence, followed by four copies of the tetracycline operator sequence, and the *CYC1* basal promoter. PCR amplification of the SAT-1-marked Tet promoter cassette was performed using 83 mer oligos which contain a 18 bp sequence complementary to the cassette and 65 bp of 5′ sequence homologous to promoter sequence corresponding to nucleotide positions −250 and −1 (relative the start codon) of the target gene facilitating precise homologous recombination between the promoter replacement fragment and the promoter of the wild-type allele following transformation. Transformants were selected on YPD medium containing nourseothricin at 400 μg ml^{-1} (Werner Bioagents). Nourseothricin-resistant colonies containing a precise promoter replacement were verified by PCR analysis. This process was repeated for all genes comprising the conditional expression strain set. A detailed description

of the GRACE methodology as applied to the target gene *KRE9* (Lussier *et al.*, 1998) is presented. Step 1: Oligonucleotide primers (99 mers) for PCR amplification of the *HIS3*-based disruption cassette were designed to contain (starting from the 3′ end of synthesized oligonucleotide) 56 nucleotides comprising a barcode element (the 3′ 18 nucleotides of which are complementary to PBSK-*HIS3*) and 43 nucleotides homologous to regions flanking the *KRE9* ORF. PCR conditions involved 5 ng PBSK-*HIS3*, 100 pmol of each primer, 150 μM dNTPs, 10 mM Tris – pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, 2.5 unit *Taq* DNA polymerase (Gibco). PCR amplification times were: 2 min at 94°C, followed by five cycles of 30 s at 94°C, 30 s at 55°C, 1 min at 72°C, an additional 25 cycles of 30 s at 94°C, 30 s at 58°C, 1 min at 72°C and 7 min at 72°C for 1 cycle. PCR-amplified *kre9Δ::CaHIS3* material was transformed into CaSS1 and histidine prototrophs were selected on YNB medium. Correct integration was confirmed by PCR analysis. Step 2: Oligonucleotide primers for PCR amplification of the conditional promoter contain 18 nucleotides complementary to the SAT-1-marked Tet promoter replacement cassette and 65 nucleotides of homologous sequence corresponding to promoter regions −270 to −205 and nucleotides 1–65 of the *KRE9* ORF. The resulting PCR product was transformed into the *kre9Δ::HIS3/KRE9* heterozygous strain produced in Step 1. Transformed cells were spread onto YPD plates and incubated overnight at 30°C, providing a pre-incubation period for expression of SAT-1 prior to replica plating onto YPD medium containing nourseothricin (400 μg ml^{-1}). Nourseothricin-resistant colonies were detected after 48 h and *kre9Δ::HIS3*/SAT-1-marked Tet-*KRE9* conditional mutants identified by PCR analysis using suitable primers which uniquely amplify the SAT-1-marked Tet-*KRE9* allele. Additional PCR genotyping was performed to confirm the presence of a *kre9Δ::HIS3* allele and the absence of a *KRE9* allele.

Phenotypic analysis of GRACE strains

Gene essentiality of all conditional mutants comprising the GRACE strain collection was evaluated in two independent methods. For example, determination of the *KRE9* GRACE strain terminal phenotype was achieved by streaking approximately 1.0×10^6 cells onto both a YNB plate and YNB plate containing 100 μg ml^{-1} tetracycline and comparing growth after 48 h at 30°C (YNB medium was chosen for phenotypic analyses so as to select for stable maintenance of both the *URA3*-marked Tet transactivator as well as the *HIS3*-marked disruption cassette). For essential genes, such as *KRE9*, no significant growth was detected on tetracycline-containing plates. Alternatively, gene essentiality was determined by streaking *KRE9* GRACE cells onto a YNB plate containing 2% glucose, 1000 μg ml^{-1} 5-flurorotic acid (5-FOA) and 100 μg ml^{-1} uridine to select for ura^- cells which have excised (via recombination between *LEU2* sequence created during targeted integration) the transactivator gene that is normally required for expression of the tetracycline promoter-regulated target gene. Again, whereas GRACE strains conditionally regulating non-essential genes demonstrate a variety of growth phenotypes under such conditions, GRACE strains for essential genes fail to grow. Applying this 5-FOA growth assay to all 574 conditional mutants displaying a 3+ shutoff growth phenotype or greater on tetracycline revealed that in 79% of the genes examined, no growth was detected on 5-FOA-containing plates, demonstrating these genes to be essential under the theoretical maximum level of repression for the Tet promoter.

Cidal versus static terminal phenotype

Quantitative evaluation of cidal versus static terminal phenotypes associated with essential GRACE strains were performed using 2×10^3 cells ml^{-1} of overnight culture inoculated into 5.0 ml of YNB liquid medium in the presence or absence of 100 μg ml^{-1} tetracycline and measuring optical density (OD_{600}) after 24 and 48 h incubation at 30°C. Typically, for essential GRACE strains, no significant increase in optical density was detected after 48 h. Distinguishing between cell death (cidal) and growth inhibitory (static) terminal phenotypes for a demonstrated essential gene was achieved by determining the percentage of viable cells [as judged by the number of colony-forming units (cfu) from an equivalent of 2×10^3 washed cells at T = 0] from the above tetracycline-treated cultures after T = 0 h, T = 20 min, T = 24 h and T = 48 h incubation. Rapid cidal targets are those displaying >90% cell death after 20 min of gene inactivation. Cidal targets are those that display reduced cell viability after 2 days versus the starting inoculum.

Bioinformatic analyses

To identify *C. albicans* homologues corresponding to *S. cerevisiae* essential genes, protein sequences of the *S. cerevisiae* essential gene set were compared with the *C. albicans* genomic sequences (www-sequence.stanford.edu/group/candida) using TBLASTN (Altschul *et al.*, 1990). For each *S. cerevisiae* essential gene, genomic DNA sequence corresponding to the best TBLASTN match was individually extracted and used to identify upstream in frame start ATG and downstream in frame stop codons to define the corresponding *C. albicans* ORF.

Scatter plot generation

The 9168 *C. albicans* ORFs predicted to encode proteins of at least 100 amino acids (www-sequence.stanford.edu/group/candida) were compared against themselves using BLASTN. A total of 2518 ORFs overlapping larger ORFs or ORFs displaying greater than 95% overlap and better than 95% identity with other ORFs were eliminated from the set to reduce dubious ORFs and allelic redundancy respectively. From this analysis, 6650 unique genes were identified to represent the *C. albicans* proteome. An independent analysis has predicted a similar gene number estimate (DeBacker *et al.*, 2001b). To construct scatter plots, protein sequences for the 6650 *C. albicans* ORFs were compared against (i) the *S. cerevisiae* genome (Sacc. Ref., http://genome-www.stanford.edu/Saccharomyces/) and (ii) the human genome (http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html), including both the annotated protein set and the Unigene set

(http://www.ncbi.nlm.nih.gov/UniGene), using either BLASTP or TBLASTN with a BLAST $P < 1 \times 10^{-4}$ as cut-off. For each *C. albicans* ORF, the *E*-values of the top match against *S. cerevisiae* or human are extracted and recorded in the scatter plot shown in Fig. 3. For simplicity, all *E*-values at $P > 1 \times 10^{-4}$ are arbitrarily defined as $P = 1 \times 10^{-4}$ and all *E*-values at $P < 1 \times 10^{-190}$ are defined as $P = 1 \times 10^{-190}$.

Immunocompetent murine model of systemic infection

An immunocompetent murine model of systemic infection was performed using male ICR mice (15 mice per group) infected via the tail vein using 1×10^6 *C. albicans* GRACE cells/mouse (Becker *et al.*, 1995; Nakayama *et al.*, 2000). Three days prior to the inoculation, five mice were given drinking water containing sucrose (5%) and tetracycline (2 mg ml^{-1}) to repress gene expression at the onset of infection. The 10 remaining mice were given drinking water with sucrose (5%) only. Five of the mice receiving sucrose water were maintained on this water throughout the infection, while five mice receiving sucrose water were switched to tetracycline-containing sucrose water at various time points after injection of *C. albicans* GRACE cells. The drinking water regimen for all mice receiving tetracycline was maintained throughout the course of the experimental infection. All *C. albicans* GRACE strains were coded prior to injection so that the laboratory workers did not know the identity of the strains. Strains were decoded at the termination of the experiment. Signs of infection (torticollis, lethargy, ataxia) were monitored and moribund mice were sacrificed. Animals were cared for according to NIH guidelines. Control experiments demonstrated that the tetracycline regimen does not affect survival of mice infected with the starting *C. albicans* strain, CaSS1, when transformed and maintaining an integrated and functional copy of *HIS3* (data not shown).

Sensitized whole-cell assays

Tetracycline concentrations which reduce the growth rate of *ERG11*, *HIS3, and ALG7* GRACE strains approximately 90% (2.5 μg ml^{-1}, 0.3 μg ml^{-1}, and 0.3 μg ml^{-1} tetracycline respectively) were determined in YNB medium (100 μl) in 96-well flat-bottom microtitre plates using initial cell inocula of ~5.0×10^4 cells of a fresh overnight culture. Cells were grown at 30°C with overnight shaking. Drug IC_{50} determinations for all strains were determined under the growth conditions described above but on medium supplemented with the appropriate tetracycline concentration. Threefold serial dilutions were performed against fluconazole, 3-aminotriazole (Sigma), tunicamycin (Sigma) and Amphotericin B (Sigma) and turbidity measured at OD_{600}. Differences in the IC_{50} value of each drug for the GRACE strains were determined by comparing IC_{50} values for each sensitized strain assayed in the presence of tetracycline versus its IC_{50} determination in the absence of tetracycline.

Acknowledgements

Sequence data for *Candida albicans* was obtained from the Stanford Genome Technology Center website at http://www-sequence.stanford.edu/group/candida. Sequencing of *Candida albicans* was accomplished with the support of the NIDR and the Burroughs Wellcome Fund. We thank A. M. Sdicu and M. Lussier for technical assistance during the early stages of this work, J. Davies for helpful advice, T. Downing for oligonucleotide synthesis, J. G. Foulkes, T. Hughes, P. T. Magee, B. B. Magee, M. Rosenberg, J. Winkelman and M. Whiteway for helpful discussions with the manuscript and P. Ocampo for manuscript assistance. T.K. and S.L. were supported by NSERC industrial fellowships 231070 and 251149 respectively.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215:** 403–410.

Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F., and Cullin, C. (1993) A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **21:** 3329–3330.

Becker, J.M., Henry, L.K., Jiang, W., and Koltin, Y. (1995) Reduced virulence of *Candida albicans* mutants effected in multi-drug resistance. *Infect Immun* **63:** 4515.

Beckerman, J., Chibana, H., Turner, J., and Magee, P.T. (2001) Single-copy IMH3 allele is sufficient to confer resistance to mycophenolic acid in *Candida albicans* and to mediate transformation of clinical Candida species. *Infect Immun* **69:** 108–114.

Braun, B.R., and Johnson, A.D. (1997) Control of filament formation in *Candida albicans* by the transcriptional repressor TUP1. *Science* **277:** 105–109.

Brown, J.S., Aufauvre-Brown, A., Brown, J., Jennings, J.M., Arst, H., Jr, and Holden, D.W. (2000) Signature-tagged and directed mutagenesis identify PABA synthetase as essential for *Aspergillus fumigatus* pathogenicity. *Mol Microbiol* **36:** 1371–1380.

Buurman, E.T., Bradley, J.D., Donnelly, C.E., Moore, J.B., Healy, J.M., Thompson, C.M., and Wobbe, C.R. (2001) Regulation of both gene expression and protein stability provides genetically assisted target evaluation (GATE) for microbial target validation. *Pharmacogenomics* **2:** 95–106.

Chandra, J., Mukherjee, P.K., Leidich, S.D., Faddoul, F.F., Hoyer, L.L., Douglas, L.J., and Ghannoum, M.A. (2001) Antifungal resistance of candidal biofilms formed on denture acrylic *in vitro*. *J Dent Res* **80:** 903–908.

DeBacker, M.D., Magee, P.T., and Pla, J. (2000) Recent developments in molecular genetics of *Candida albicans*. *Annu Rev Microbiol* **54:** 463–498.

DeBacker, M.D., Nelissen, B., Logghe, M., Viaene, J., Loonen, I., Vandoninck, S., *et al.* (2001a) An antisense-based functional genomics approach for identification of genes critical for growth of *Candida albicans*. *Nature Biotechnol* **19:** 235–241.

DeBacker, M.D., Ilyina, T., Ma, X.J., Vandoninck, S., Luyten, W.H., and Vanden Bossche, H. (2001b) Genomic profiling of the response of *Candida albicans* to itraconazole treatment using a DNA microarray. *Antimicrob Agents Chemother* **45:** 1660–1670.

Devasahayam, G., Chaturvedi, V., and Hanes, S.D. (2002) The Ess1 prolyl isomerase is required for growth and mor-

phogenetic switching in *Candida albicans*. *Genetics* **160:** 37–48.

DiDomenico, B. (1999) Novel antifungal drugs. *Curr Opin Microbiol* **2:** 509–515.

Douglas, C.M., D'Ippolito, J.A., Shei, G.J., Meinz, M., Onishi, J., Marrinan, J.A., *et al.* (1997) Identification of the *FKS1* gene of *Candida albicans* as the essential target of 1,3-beta-D-glucan synthase inhibitors. *Antimicrob Agents Chemother* **41:** 2471–2479.

Enloe, B., Diamond, A., and Mitchell, A.P. (2000) A single-transformation gene function test in diploid *Candida albicans*. *J Bacteriol* **182:** 5730–5736.

Fonzi, W.A., and Irwin, M.Y. (1993) Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* **134:** 717–728.

Forsyth, R.A., Haselbeck, R.J., Ohlsen, K.L., Yamamoto, R.T., Howard, H., Trawick, J.D., *et al.* (2002) A genome-wide strategy for the identification of essential genes in *Staphylococcus aureus*. *Mol Microbiol* **43:** 1387–1400.

Gari, E., Piedrafita, L., Aldea, M., and Herrero, E. (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast* **13:** 837–848.

Giaever, G., Shoemaker, D.D., Jones, T.W., Liang, H., Winzeler, E.A., Astromoff, A., and Davis, R.W. (1999) Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nat Genet* **21:** 278–283.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., *et al.* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418:** 387–391.

Goffeau, A.,, Alert, R., Agostini-carbone, M.L., Ahmed, A., Aigle, M., Alberghina, L., *et al.* (1997) The Yeast Genome Directory. *Nature* **387:** 1–105.

Gow, N.A., Robbins, P.W., Lester, J.W., Brown, A.J., Fonzi, W.A., Chapman, T., and Kinsman, O.S. (1994) A hyphal-specific chitin synthase gene (CHS2) is not essential for growth, dimorphism, or virulence of *Candida albicans*. *Proc Natl Acad Sci USA* **91:** 6216–6220.

Guthrie, C., and Fink, G.R. (1991) Guide to yeast genetics and molecular biology. *Methods Enzymol* **194:** 302–318.

Hartman, J.L., Garvik, B., and Hartwell, L. (2001) Principles for the buffering of genetic variation. *Science* **291:** 1001–1004.

Heckman, D.S., Geiser, D.M., Eidell, B.R., Stauffer, R. L., Kardos, N.L., and Hedges, S.B. (2001) Molecular evidence for the early colonization of land by fungi and plants. *Science* **293:** 1129–1133.

Hughes, T.R., Marton, M.J., Jones, A.R., Roberts, C.J., Stoughton, R., Armour, C.D., *et al.* (2000) Functional discovery via a compendium of expression profiles. *Cell* **102:** 109–126.

Kelly, R., Card, D., Register, E., Mazur, P., Kelly, T., Tanaka, K.I., *et al.* (2000) Geranylgeranyltransferase I of *Candida albicans*: Null mutants or enzyme inhibitors produce unexpected phenotypes. *J Bacteriol* **182:** 704–713.

Lussier, M., Sdicu, A.M., Shahinian, S., and Bussey, H. (1998) The *Candida albicans KRE9* gene is required for cell wall β-1,6-glucan synthesis and is essential for growth on glucose. *Proc Natl Acad Sci USA* **95:** 9825–9830.

Madhani, H.D., and Fink, G.R. (1998) The control of filamentous differentiation and virulence in fungi. *Trends Cell Biol* **8:** 348–353.

Magee, P.T. (1993) Variations in chromosome size and organization in *Candida albicans* and *Candida stellatoidea*. *Trends Microbiol* **1:** 338–342.

Malpertuy, A., Tekaia, F., Casaregola, S., Aigle, M., Artiguenave, F., Blandin, G., *et al.* (2000) Genomic exploration of the Hemiascomycetous Yeast: 19. Ascomycetes-specific genes. *FEBS Microbiol lett* **487:** 113–121.

Mio, T., Yamada-Okabe, T., Yabe, T., Nakajima, T., Arisawa, M., and Yamada-Okabe, H. (1997) Isolation of the *Candida albicans* homologs of *Saccharomyces cerevisiae KRE6* and *SKN1*: Expression and physiological function. *J Bacteriol* **179:** 2363–2372.

Mitchell, A.P. (1998) Dimorphism and virulence in *Candida albicans*. *Curr Opin Microbiol* **1:** 687–692.

Nagahashi, S., Mio, T., Ono, N., Yamada-Okabe, T., Arisawa, M., Bussey, H., and Yamada-Okabe, H. (1998) Isolation of *CaSLN1* and *CaNIK1*, the genes for osmosensing histidine kinase homologues, from the pathogenic fungus *Candida albicans*. *Microbiol* **144:** 425–432.

Nakayama, H., Mio, T., Nagahashi, S., Kokado, M., Arisawa, M., and Aoki, Y. (2000) Tetracycline-regulatable system to tightly control gene expression in the pathogenic fungus *Candida albicans*. *Infect Immun* **68:** 6712–6719.

Odds, F.C., Gow, N.A., and Brown, A.J. (2001) Fungal virulence studies come of age. *Genome Biol* **2:** 1009–1032.

Rustchenko, E.P., Howard, D.H., and Sherman, F. (1994) Chromosomal alterations of *Candida albicans* are associated with the gain and loss of assimilating functions. *J Bacteriol* **176:** 3231–3241.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Santos, M.A., and Tuite, M.F. (1995) The CUG codon is decoded *in vivo* as serine and not leucine in *Candida albicans*. *Nucleic Acids Res* **23:** 1481–1486.

Shoemaker, D.D., Lashkari, D.A., Morris, D., Mittmann, M., and Davis, R.W. (1996) Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat Genet* **14:** 450–456.

Sobel, J.D. (2000) Practice guidelines for the treatment of fungal infections. *Clin Infectious Dis* **30:** 652.

Stoldt, V.R., Sonneborn, A., Leuker, C.E., and Ernst, J.F. (1997) Efg1p, an essential regulator of morphogenesis of the human pathogen *Candida albicans*, is a member of a conserved class of bHLH proteins regulating morphogenetic processes in fungi. *EMBO* **16:** 1982–1991.

Warit, S., Zhang, N., Short, A., Walmsley, R.M., Oliver, S.G., and Stateva, L.I. (2000) Glycosylation deficiency phenotypes resulting from depletion of GDP-mannose pyrophosphorylase in two yeast species. *Mol Microbiol* **36:** 1156–1166.

Wilson, R.B., Davis, D., and Mitchell, A.P. (1999) Rapid hypothesis testing with *Candida albicans* through gene disruption with short homology regions. *J Bacteriol* **181:** 1868–1874.

Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., *et al.* (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285:** 793–972.